(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[7]: **H04N 7/24**

(21) International Application Number:
PCT/KR2003/000101

(22) International Filing Date: 16 January 2003 (16.01.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10-2002-0067299 31 October 2002 (31.10.2002) KR
10-2003-0002074 13 January 2003 (13.01.2003) KR

(71) Applicant (for all designated States except US): ELEC-TRONICS AND TELECOMMUNICATIONS RE-SEARCH INSTITUTE [KR/KR]; 161, Gajeong-dong, Yuseong-gu, Daejon 305-350 (KR).

(72) Inventors; and
(75) Inventors/Applicants (for US only): PARK, Sung-Hee [KR/KR]; #105-104, Daerim Doore Apt., Sinsung-dong, Yuseong-gu, Daejon 305-720 (KR). PARK, Soo-Jun [KR/KR]; #203-501, Hyundai Apt., 477 Apgujeong-dong, Gangnam-gu, Seoul 135-903 (KR). JANG, Myung-Gil [KR/KR]; #103-504, Dasol Apt., Gung-dong, Yuseong-gu, Daejon 305-335 (KR). PARK, Sang-Kyu [KR/KR]; #105-1305, Hanmaeul Apt., Songgang-dong, Yuseong-gu,

Daejon 305-756 (KR). WON, Chee-Sun [KR/KR]; 3-26 Pil-dong, Jung-gu, Seoul 100-866 (KR).

(74) Agent: SHINSUNG PATENT FIRM; Haecheon Bldg., 741-40, Yeoksam 1-dong, Kangnam-ku, Seoul 135-924 (KR).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NON-LINEAR QUANTIZATION AND SIMILARITY MATCHING METHODS FOR RETRIEVING VIDEO SEQUENCE HAVING A SET OF IMAGE FRAMES
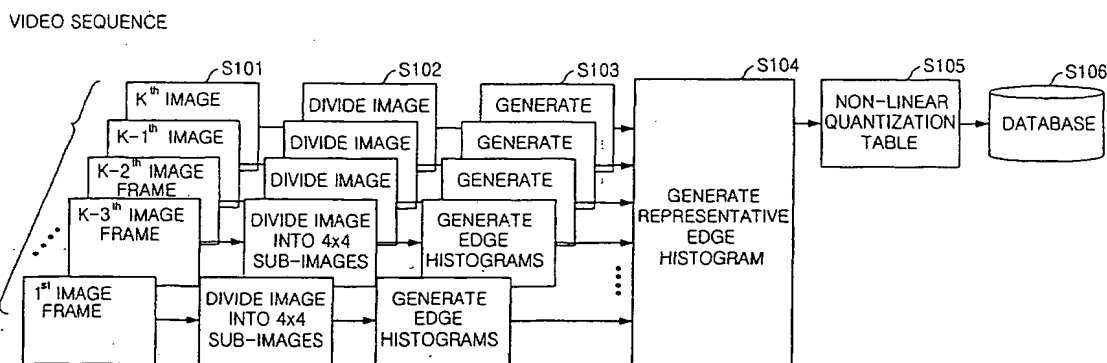
VIDEO SEQUENCE

(57) Abstract: A method for constructing a database having digital video data information representing a plurality of video sequence is disclosed. The method includes the steps of: a) partitioning each image frame of each video sequence into L number of sub images; b) generating L number of edge histograms for each image frame; c) normalizing the edge histogram bins to thereby generate M number of normalized edge histogram bins; d) calculating M representative edge histogram bins in order to generate L number of representative edge histograms based on the normalized edge histogram bins; and e) non-linearly quantizing the representative edge histogram bins to generate M number of quantization index values for the each representative edge histogram, to be stored in the database.

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

NON-LINEAR QUANTIZATION AND SIMILARITY MATCHING METHODS FOR RETRIEVING VIDEO SEQUENCE HAVING A SET OF IMAGE FRAMES

Field of the Invention

5

This invention relates to a method for retrieving image data; and, more particularly, to a method for constructing a bit-representation of an edge histogram descriptor having reduced bits for a video sequence

10 including a set of image frames and a method for retrieving a video sequence by using the information effectively extracted from the coded representation of the edge histogram descriptor.

15   ## Description of the prior Art

The Joint Photographic Experts Group (JPEG) is the international standard for still images and the Moving Picture Experts Group-1 (MPEG-1), 2 are for moving pictures.

20 Regarding compressed image information, feature information for each image is to be extracted for applications such as extracting a key frame, image search, browsing or the like.

To extract the feature information, brightness or color histograms are widely used. The brightness and color

25 histograms, respectively, represent relative frequency of brightness and color (red, green or blue) in an image. Especially, various methods of histogram comparison have been recently proposed for searching still images or digital video data that are stored digitally. As the

30 histograms get used for image search and shot boundary detection, it is believed that conventional histograms are to be improved. That is, it is required to adopt a histogram descriptor such as an edge histogram, which can represent the image content more efficiently. Also, the

35 binary representation of the descriptor should be compact and the computational complexity for the similarity matching should be low.

A method of employing color histograms and edge maps

for a shot boundary detection is disclosed in U.S. Patent
No. 5,805,733 entitled "Method and system for detecting
scenes and summarizing video sequences". Though the method
is effective in that color information considering a human
5   visual system is extracted, it does not include extraction
of brightness information.

A method according to which color information is
received and then indexing is performed by measuring
similarity of images using histogram intersection technique
10 is disclosed in an article by M. J. Swain, et al., "Color
Indexing", *International Journal of Computer Vision, Vol.*
7-1, pp.11-32, 1991. However, this method does not use
brightness and edge information and thus accuracy is not
guaranteed enough. Also, since the histograms are
15 generated using a discrete quantization technique in the
conventional methods, relatively a large number of
histogram bins are necessary to make equal performance.
Consequently, inefficiency in storage and similarity
measurement is caused. In addition, because a feature
20 extraction is performed in terms of pixel conventionally,
there is a problem that feature information is
restrictively generated.

In the meantime, as the histogram is widely used for
image searching, etc., recently, the efficient way of
25 storing histogram information is required. In other words,
a histogram bin value is stored in a storage field of fixed
size by normalization using linear quantization according
to a conventional way to store the histogram. Consequently,
this method of the linear quantization to the histogram
30 storage causes a problem as quantity of bits is increased.

The International Organization for Standardization/
International Electrotechnical Commission Joint Technical
Committee 1 (ISO/IEC JTC1) establishes international
standards for a content based multimedia retrieval
35 technique related to the MPEG-7. A content based
multimedia includes a moving picture and still images such
as a digital video data. The digital video data i.e., a
video sequence contains a number of image frames of at

least one moving object. For retrieving the video sequence, a motion descriptor for moving object is extracted from the image frames, wherein the motion descriptors contain motion information of the moving objects in the image frames.

5   After extracting the motion descriptor, likelihood between an inquiry video sequence and motion descriptors of video sequence stored in the database is computed. Finally, according to a computed likelihood, a desired video sequence is retrieved.

10   Generally, a motion trajectory descriptor is widely used as a motion descriptor in the content based multimedia retrieval technique. The motion trajectory descriptor contains information of motion trajectories of moving objects in image frames of the video sequence and the

15 motion trajectories of the moving objects by using a parametric equation based on locations of objects and a speed of moving objects. In a conventional method by using the motion trajectory descriptor, it is hard to represent a "texture video sequence", which has a number of moving

20 objects, such as a video data containing images of a firework or a waterfall. That is, in the texture video sequence, there are great number of moving objects to be represented by using the motion trajectory descriptors. As a result, there is a great computation burden for

25 extracting the great number of the motion trajectory descriptors for great number of moving objects.

       Therefore, for effectively retrieval digital video data including the texture video sequence, a new digital video data retrieval method and enhanced description scheme

30 have been demanded.

       Summary of the Invention

       It is an object of the present invention to provide a

35 method for constructing a database having image information representing a plurality of video sequence with reduced bits to be stored in the database.

       It is another object of the present invention to

provide a method for retrieving a corresponding video sequence in response to a query video sequence based on a database with a high retrieval speed and accuracy.

It is still another object of the present invention to provide a method for retrieving a corresponding video sequence including a texture video in response to a query video sequence based on a database with a high retrieval speed and accuracy.

In accordance with one aspect of the present invention, there is provided a method for constructing a database having digital video data information representing a plurality of video sequence, each video sequence having a set of image frames of the digital video data, the method including the steps of: a) partitioning each image frame of each video sequence into L number of sub images, wherein each sub-image is further partitioned into S x T number of image-blocks, L, S and T being positive integers; b) assigning one of 5 number of reference edges to each image-block to thereby generate L number of edge histograms for each image frame, wherein the edge histogram includes the M edge histogram bins and the reference edges include 4 number of directional edges and a non-directional edge; c) normalizing the edge histogram bins contained in each edge histogram by S x T to thereby generate M number of normalized edge histogram bins for the each image frame;

d) calculating M representative edge histogram bins of the each video sequence in order to generate L number of representative edge histograms of each video sequence based on the normalized edge histogram bins of the each image frames; and e) non-linearly quantizing the representative edge histogram bins to generate M number of quantization index values as a second image descriptor for the each representative edge histogram to be stored in the database.

In accordance with another aspect of the present invention, there is provided a method for retrieving a corresponding video sequence having a set of image frames of the digital video data in response to a query video sequence based on a database, the method including the

steps of: a) calculating L number of representative edge histograms of the query video sequence as an image descriptor for the query video sequence, wherein each representative edge histogram represents a representative
5 spatial distribution of 5 number of reference edges in sub-images of image frames in the query video sequence, wherein the reference edges includes 4 number of directional edges and a non-directional edge; b) extracting a plurality of image descriptors for video sequences based on digital
10 video data information from the database, wherein each image descriptor for the each video sequence includes L number of representative edge histogram bins for said each video sequence; c) comparing the image descriptor for the query video sequence to said each image descriptor for each
15 video sequences to generate a comparison result; and d) retrieving at least one target video sequence similar to the query video sequence based on the comparison results.

In accordance with still another aspect of the present invention, there is provided a method for extracting an
20 image descriptor for a video sequence, each having a plurality of image frames of the digital video data, the method comprising the steps of: a) selecting a one of the image frames for a target video sequence as a target image frame; b) calculating Lx5 number of normalized edge
25 histogram bins to generate L number of edge histograms of the target image frame, wherein the each edge histogram has 5 number of normalized edge histogram bins and represents a spatial distribution of 5 number of reference edges in a sub-image and L is a positive integer, wherein the
30 reference edges include 4 number of directional edges and a non-directional edge; c) selecting a next image frame as a target image frame; d) repeating steps b) and c) until L number of edge histograms of all image frames are calculated; e) calculating a representative edge histogram
35 having LX5 number of normalized edge histogram bins for the video sequence based on the L number of edge histograms of each image frame; f) non-linearly quantizing the Lx5 number of normalized edge histogram bins of the

representative edge histogram to generate Lx5 number of quantization index values for the digital video data as the image descriptor for the video sequence; and g) storing the Lx5 number of quantization index values to the database.

Brief Description of the Drawings

Other objects and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, in which:

Fig. 1A is a block diagram illustrating a parallel process for constructing a database having a plurality of image descriptors for corresponding video sequences in accordance with one embodiment of the present invention;

Fig. 1B is a flowchart explaining a serial process for constructing a database having a plurality of image descriptors for corresponding video sequences in accordance with an other embodiment of the present invention;

Fig. 1C is a flowchart explaining a serial process for constructing a database having a plurality of image descriptors for corresponding video sequence in accordance with another embodiment of the present invention;

Fig. 2 shows an explanatory diagram depicting an image having 16 sub-images to be represented by image descriptors;

Figs. 3A to 3E illustrate 5-types of edges to be used for an edge determination process in accordance with the present invention;

Fig. 4 is an explanatory diagram demonstrating an image-block partitioned into 4 sub-blocks each of which filter coefficients are assigned to;

Figs. 5A to 5E show image-blocks, wherein sub-blocks of each image-block are provided with corresponding filer coefficients for 5 edges;

Fig. 6 is an explanatory diagram illustrating an array of 80 edge histogram bins corresponding to each image frame; and

Fig. 7 is a diagram showing a process for retrieving a

desired video sequence in response to an input of query video sequence in accordance with the present invention.

Detailed Description of the Preferred Embodiments

Hereinafter, preferred embodiments in accordance with the present invention will be described in detail referring to the accompanying drawings.

Referring to Fig. 1A, there is shown a block diagram illustrating a parallel process for constructing a plurality of image descriptors for corresponding video sequences in accordance with one embodiment of the present invention.

As shown, the target video sequence includes a number of the image frames and edge histograms of each image frame are simultaneously generated.

At the processing block S101, k number of image frames are imputed to a processing block S102. At the processing block S102, each image frame is divided into NxN, e.g., 4x4 sub-images, wherein N is a positive integer. The sub-images for each image frame are then coupled to a processing block S103 for generating edge histograms of each image frame of the video sequence. That is, an edge histogram for each sub-image is obtained by using a multiplicity of edges and then 80 normalized local edge histogram bins for each image frame are coupled to the processing block S104.

At the processing block S104, representative edge histograms of the target video sequence are computed as a first image descriptor by calculating 80 representative edge histogram bins based on 80 normalized edge histogram bins of each image frames in the video sequence.

Each of the representative edge histogram bins may be one of a mean value or a median value of corresponding normalized edge histogram bins of each image frame. Also, each of the representative edge histogram bins may be one of corresponding normalized edge histogram bins of each image frame by selecting an intersection value or a key

value among the corresponding normalized edge histogram bins.

In a meantime, after calculating the representative edge histogram, other statistical values representing variation of objects can be also used for retrieving a desired video sequence together with the representative edge histograms. The other statistical value includes a variance representing the difference between two or more image frames.

At the processing block S105, the representative edge histograms are non-linearly quantized to thereby generate a corresponding second image descriptor which is, e.g., a group of quantization index values.

Thereafter, the second image descriptor for the target video sequence is inputted and stored in a database S106. The above process is performed by using a number of video sequence to be stored in the database.

Referring to Fig. 1B, there is shown a flowchart explaining a serial process for constructing database having a plurality of image descriptors for corresponding video sequences, each video sequence including a set of image frames of digital video data, in accordance with the present invention.

As described hereinbefore, the video sequence includes a number of the image frames and edge histograms of each image frame are serially generated in order to obtain the representative edge histograms.

At step of S110, one of image frames in the video sequence is selected as a target image frame. At step S111, the selected image frame is divided into NxN, e.g., 4x4 sub-images, wherein N is a positive integer. Edge histograms are extracted from the sub-image at step of S112. At step of S113, it is determined whether or not the edge histograms of all sub-images are generated. If the edge histograms of all sub-images are not generated then a next sub-image is selected at step S114 and the edge histogram of the next sub-image is generated at step 112. Otherwise, if the edge histograms of all sub-images are generated then

integer number k is increased by one for selecting next
image frame in the video sequence at step of S115. After
increasing the k, it is determined whether all image frames
of the video sequence are selected. If all image frames of
5   the video sequence are not selected then a next image frame
is selected as a newly selected target frame to be
processed and the above mentioned steps S110 to S115 are
repeated. That is, the edge histograms for each sub-image
is obtained by using a multiplicity of edges contained in
10 each sub-images and then 80 normalized local edge histogram
bins for each image frame.

After generating all edge histograms of all image
frames in the video sequence, representative edge
histograms are generated as a first image descriptor at
15 step 117 by calculating 80 representative edge histogram
bins based on 80 normalized edge histogram bins of each
image frames. Each of the representative edge histogram
bins may be one of a mean value or a median value of
corresponding normalized local edge histogram bins of all
20 image frame. Also, each of the representative edge
histogram bins may be one of corresponding normalized local
edge histogram bins of all image frames by selecting an
intersection value or a key value among the normalized
edge histogram bins.

25 The representative edge histograms are non-linearly
quantized to thereby generate a corresponding second image
descriptor which is, e.g., a group of quantization index
values at step S118. Thereafter, the second image
descriptor for the video sequence is inputted and stored in
30 a database. The above process is repeated until all video
sequences to be stored in the database are processed.

Fig. 1C is a flowchart explaining a serial process for
constructing a database having a plurality of image
descriptors for corresponding video sequence having image
35 frames of digital video data in accordance with another
preferred embodiment of the present invention. Referring
to Fig. 1C, the serial process for constructing a database
is identical to the flowchart in Fig. 1B except a step 119.

Therefore, for the sake of convenience, detailed explanation of steps S110 to S117 is omitted.

After generating the representative edge histogram at step S117, a variation value representing the difference
5 between two or more image frames is computed at step S119. The variation value can be obtained by calculating a variation of the edge histogram of each image frame in the digital video data. The variation values representing variation of objects can be also used for retrieving the
10 desired video sequence. A variance or a standard deviation can be the variation values. Together with the representative edge histograms, the variance may be used for retrieving the digital video data in detail.

Referring to Figs. 2 to 6, there are shown explanatory
15 diagrams for illustrating a process for obtaining the first image descriptor described referring to Fig. 1.

As shown in Fig. 2, in order to obtain a corresponding edge histograms of each image frame in a video sequence, an inputted image frame of digital 200 is divided into 4×4
20 non-overlapping sub-images to thereby form 16 number of rectangle-shaped sub-images 211 to 226. Each of the sub-images contains a plurality of pixels.

In order to extract the edge histogram, each sub-image is then divided into M × T non-overlapping square-shaped
25 image-blocks wherein the size of the image-block depends on the size of the image. Each image-block is used in an edge determination process, in which an image-block is described by using one of edges.

In accordance with one embodiment of the present
30 invention, as shown in Figs. 3A to 3E, the edge determination process is provided with five edges, one of which is selected for an image-block. The edges can include various types of directional edges, preferably, vertical, horizontal, 45 degree and 135 degree edges 301 to
35 307; and a non-directional edge 309 including at least one edge of undesignated direction.

In order to generate an edge histogram for a sub-image, it is necessary to detect an edge feature from an image-

block. That is, the edge determination process is
performed in order to determine which one of edges can be
assigned to an image- block. The extraction can be
performed using a method applying digital filters in the
5   spatial domain.

In the edge determination process, as shown in Fig. 4,
an image-block is partitioned into 4 sub-blocks. That is,
as shown, a reference numeral 400 denotes the image-block
and reference numerals 411, 413, 415 and 417 denote sub-
10   blocks, respectively. The sub-blocks are also labeled 0, 1,
2 and 3 for the image-block 400, wherein a corresponding
filter coefficient is assigned to each sub-block to be able
to obtain a set of edge magnitudes.

In accordance with one embodiment of the present
15   invention, each image-block 400 is partitioned into 2×2
sub-blocks each of which is labeled 0, 1, 2 or 3.

For each image block, a set of five edge magnitudes
corresponding to five types of edges are obtained by using
following equations:

20
$$m_v(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_v(k) \right|$$ Eq. 1

$$m_h(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_h(k) \right|$$ Eq. 2

$$m_{d-45}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-45}(k) \right|$$ Eq. 3

$$m_{d-135}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-135}(k) \right|$$ Eq. 4

$$m_{nd}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{nd}(k) \right|$$ Eq. 5

25   where $m_v(i,j)$, $m_h(i,j)$, $m_{d-45}(i,j)$, $m_{d-135}(i,j)$ and $m_{nd}(i,j)$,
respectively, denote vertical, horizontal, 45 degree, 135
degree and non-directional edge magnitudes for a (i,j) [th]
image-block; $a_k(i,j)$ denotes an average gray level for a sub-
block labeled k in the (i,j) [th] image-block and $f_v(k)$, $f_h(k)$,
30   $f_{d-45}(k)$, $f_{d-135}(k)$ and $f_{nd}(k)$ denote, respectively, filter
coefficients for the vertical, horizontal, 45 degree, 135

degree and non-directional edges where k = 0, 1, 2 and 3, each representing a number labeling each sub-block.

Referring to Figs. 5A to 5E, there are shown filter coefficients for each edge.

As shown, reference numerals 501, 503, 505, 507 and 509, respectively, show respective filter coefficients for the vertical, horizontal, 45 degree, 135 degree and non-directional edges. Each image-block can be represented by using a selected one among five edge magnitudes, wherein each magnitude is calculated for respective edges.

In order to determine an edge corresponding to an image-block, the five edge magnitudes obtained by using the above equations are compared each other. According to the comparison, the image-block is expressed by one edge having a maximum edge magnitude among them, where the maximum edge magnitude should be also greater than a predetermined threshold value. If the maximum edge magnitude is less than the predetermined threshold value, it is determined that the image-block contains no edge.

When the selected edge for the image-block is determined as a result of the edge magnitude comparison, a corresponding edge histogram bin for the sub-image is increased by 1. There are five types of edge histogram bins, i.e., vertical, horizontal, 45 degree, 135 degree and non-directional edge histogram bins. The five edge histogram bins are components for representing the edge histogram. The detection of corresponding edges for all of the image-blocks included in a sub-image is performed and then an edge histogram bin corresponding to each detected edge is increased by 1 to thereby generate an edge histogram, so called, a local edge histogram, for the sub-image. The detection of edges and generation of edge histogram are performed for all of 16 sub-images.

The local edge histogram represents the distribution of 5 types of edges in a sub-image, i.e., an edge histogram for a sub-image. Since the number of sub-images is fixed to 16 and each sub-image is assigned 5 edge histogram bins, 80 edge histogram bins are needed to generate corresponding

local edge histograms for all of 16 sub-images. That is, semantics of each bin of BinCounts is defined as shown in following Table 1:

5   Table 1

Edge histogram bin	Semantics
BinCounts[0]	Vertical edges in sub-image (0,0)
BinCounts[1]	Horizontal edges in sub-image (0,0)
BinCounts[2]	45 degree edges in sub-image (0,0)
BinCounts[3]	135 degree edges in sub-image (0,0)
BinCounts[4]	Non-directional edges in sub-image (0,0)
BinCounts[5]	Vertical edges in sub-image (0,1)
.	.
.	.
BinCounts[74]	Non-directional edges in sub-image (3,2)
BinCounts[75]	Vertical edges in sub-image (3,3)
BinCounts[76]	Horizontal edges in sub-image (3,3)
BinCounts[77]	45 degree edges in sub-image (3,3)
BinCounts[78]	135 degree edges in sub-image (3,3)
BinCounts[79]	Non-directional edges in sub-image (3,3)

where BinCount[0], BinCount[1] . . . BinCount[79] represent coded bits for the edge histogram descriptor.

Referring to Fig. 6, there is shown exemplary arrays
10   of 80 edge histogram bins corresponding to each image frame in a video sequence.

For example, a edge histogram for a sub-image 211 at (0,0) of the image 200 shown in Fig. 2 includes vertical, horizontal, 45 degree, 135 degree and non-directional edge
15   histogram bins 600, 601, 602, 603 and 604 of 1st image frame (which is referred to BIN COUNT[1.0], BIN COUNT[1.1], BIN COUNT[1.3] (not shown) and BIN COUNT[1.4] (not shown) as shown in Fig. 6). In the same way, a local edge histogram for a sub-image 212 at (0,1) in Fig. 2 includes 5 edge
20   histogram bins 605, 606, 607, 608 and 609 (which is also

referred to BIN COUNT[1.5], BIN COUNT[1.6], BIN COUNT[1.7] and BIN COUNT[1.9] (not shown)) in the same sequence as that of bins for the sub-image 211. Consequently, total 80 edge histogram bins are needed to generate respective 16

5   edge histograms for all of the 16 sub-images, wherein the total 80 bins are calculated by multiplying 5 edge histogram bins by 16 sub-images.

In order to obtain edge histograms of each image frame of the video sequence, each edge histogram bin in a local

10   edge histogram for a sub-image is normalized by dividing each bin by the total number of the image-blocks included in the sub-image. Thereby, each edge histogram bin for the local edge histogram has a bin value ranging from 0 to 1.

After computing all of edge histograms of each image

15   frame in the digital video data, representative edge histograms of the video sequence are computed as a first image descriptor by calculating 80 representative edge histogram bins based on 80 normalized local edge histogram bins of image frame in the video sequence.

20   Each of the representative edge histogram bins may be one of a mean value or a median value of corresponding normalized edge histogram bins in all image frames. Also, each of the representative edge histogram bins may be one of corresponding normalized edge histogram bins of all

25   image frames by selecting an intersection value or a key value among the local edge histogram bins located at the same position.

For example, if the mean value is used for calculating the representative edge histograms, then the

30   representative edge histograms are calculated as following. Referring to Fig. 6, corresponding edge histogram bins located at the same position for each image frame are added and divided by the number of frame in the video sequence to generate representative edge histogram bins. For example,

35   corresponding edge histogram bins BIN COUNT [k.0], BIN COUNT[k-1,0],… BIN COUNT[1.0] are added and divided by the number of frame to generate the representative edge histogram bin BIN COUNT[0]. All other edge histogram bins

are also added and divided by the number of frame in the digital video data to generate the representative edge histogram bins such as BIN COUNT[0], BIN COUNT[1] ... BIN COUNT[79]. After calculating all representative edge
5 histogram bins, the representative edge histograms of the video sequence are stored as the first image descriptor of the video sequence as mentioned above.

The normalized bin values of representative edge histogram in the digital video data are then coupled to the
10 processing block S105 shown in Fig. 1A as the representative histogram bins. At the processing block S105, the representative histogram is non-linearly quantized by using a number of quantization tables.

That is, in order to obtain a second image descriptor,
15 the normalized bin values are quantized for obtaining binary representations thereof. The quantization should be performed for the normalized 80 bin values of the representative edge histograms. In this case, the normalized bin values are non-linearly quantized to be able
20 to minimize overall number of bits for the binary representations. The above process is performed for all video sequence to be stored in a database.

As a result, a group of quantization index values is obtained as the second image descriptor. The non-linear
25 quantization is performed using a non-linear quantizer designed with, e.g., a Lloyd-Max algorithm in accordance with one embodiment of the present invention.

In order to perform the quantization, five non-linear quantization tables for each of vertical edge, horizontal
30 edge, 45 degree edge, 135 degree edge and non-directional edge histogram bins are used therein, which can be represented as following listed Tables 2 to 6:

Table 2: quantization table for the vertical edge histogram bin

Index(3bits/bin)	Range	Representative value
0	0.0000000 ~ 0.0343910	0.010867

1	0.0343910 ~ 0.0787205	0.057915
2	0.0787205 ~ 0.1221875	0.099526
3	0.1221875 ~ 0.1702110	0.144849
4	0.1702110 ~ 0.2280385	0.195573
5	0.2280385 ~ 0.3092675	0.260504
6	0.3092675 ~ 0.4440795	0.358031
7	0.4440795 ~ 1.0000000	0.530128

Table 3: quantization table for the horizontal edge histogram bin

5

Index(3bits/bin)	Range	Representative value
0	0.0000000 ~ 0.0411000	0.012266
1	0.0411000 ~ 0.0979065	0.069934
2	0.0979065 ~ 0.1540930	0.125879
3	0.1540930 ~ 0.2128515	0.182307
4	0.2128515 ~ 0.2789795	0.243396
5	0.2789795 ~ 0.3631455	0.314563
6	0.3631455 ~ 0.4880235	0.411728
7	0.4880235 ~ 1.0000000	0.564319

Table 4: quantization table for the 45 degree edge histogram bin

Index(3bits/bin)	Range	Representative

		value
0	0.0000000 ~ 0.0150225	0.004193
1	0.0150255 ~ 0.0363560	0.025852
2	0.0363560 ~ 0.0576895	0.046860
3	0.0576895 ~ 0.0809025	0.068519
4	0.0809025 ~ 0.1083880	0.093286
5	0.1083880 ~ 0.1424975	0.123490
6	0.1424975 ~ 0.1952325	0.161505
7	0.1952325 ~ 1.0000000	0.228960

Table 5: quantization table for the 135 degree edge histogram bin

Index(3bits/bin)	Range	Representative value
0	0.0000000 ~ 0.0150490	0.004174
1	0.0150490 ~ 0.0360780	0.025924
2	0.0360780 ~ 0.0566975	0.046232
3	0.0566975 ~ 0.0784090	0.067163
4	0.0784090 ~ 0.1025230	0.089655
5	0.1025230 ~ 0.1336475	0.115391
6	0.1336475 ~ 0.1848245	0.151904
7	0.1848245 ~ 1.0000000	0.217745

5

Table 6: quantization table for the non-directional edge histogram bin

Index(3bits/bin)	Range	Representative value
0	0.0000000 ~ 0.0292225	0.006778
1	0.0292225 ~ 0.0801585	0.051667
2	0.0801585 ~	0.108650

		0.1374535	
3	0.1374535 ~ 0.1952415	0.166257
4	0.1952415 ~ 0.2549585	0.224226
5	0.2549585 ~ 0.3210330	0.285691
6	0.3210330 ~ 0.4036735	0.356375
7	0.4036735 ~ 1.0000000	0.450972

where the optimal number of bits per bin is fixed to 3 in order to have 8 quantization levels in the above quantization tables in accordance with present invention.

5 The second image descriptor is then stored in the database S106 and will be retrieved in response to an input of a query image.

Fig. 7 is a diagram illustrating a method for retrieving a desired video sequence in response to an input

10 of query video sequence in accordance with a preferred embodiment of the present invention.

If a query video sequence is received, the query video sequence is processed in the same manner of the processing blocks S101 and S103 shown in Fig. 1A. That is, edge

15 histograms of each image frame for the query video sequence are obtained by using the above same manner and includes normalized edge histogram bins for the query video sequence.

Thereafter, local edge histograms of each image frame in the video sequence, representative edge histograms of

20 the video sequence, global edge histogram and semi-global histograms for the query video sequence are generated based on the normalized edge histogram bins as an image descriptor. The global edge histogram represents the edge distribution for the whole image space. The global edge

25 histogram and the semi-global histograms will be described hereinafter in more detail.

On the other hand, referring to Fig. 7, there is shown a method for retrieving a desired digital video data in response to an input of a query video sequence by using a number of non-linear inverse quantization tables in accordance with a preferred embodiment of the present invention, wherein the non-linear inverse tables can be of Tables 2, 3, 4, 5 and 6 as described above.

When a query video sequence is inputted, the same process as that in the processing block S101, i.e., the image division process is performed at a processing block S701.

At a processing block S702, the same process as that in the processing block S103, i.e., the edge histogram generation of each image frame is performed.

At a processing block S703, the representative edge histogram of the video sequence is generated according to the edge histograms of each image frame in the video sequence.

After the calculating the representative edge histograms of the video sequence, a non-linear quantization process is performed same as processing block S105 in Fig. 1A.

In order to achieve a high retrieval performance, a global edge histogram and semi-global edge histograms for the query video sequence can be further generated after non-linear inverse quantization process of the representative edge histogram S704, based on the representative edge histogram bins that are generated at the processing block S703.

For a data matching process, a number of second image descriptors for each video sequence are retrieved sequentially from the pre-established database S107. For a stored target video sequence, a group of quantization index values are retrieved and coupled to the non-linear inverse-quantization tables S704. Through the use of the non-linear inverse-quantization tables, the quantization index values are then converted into normalized edge histogram bins for the retrieved video sequence.

At a processing block S705, the representative edge histograms of the query video sequence and the retrieved video sequence is compared for selecting a desired video sequence.

5   For minutely matching the retrieved video sequence and the query video sequence, a global edge histogram and semi-global edge histograms can be used. For example, the data matching process block S705 is explained by using the global edge histogram and the semi-global edge histograms.

10   The normalized edge histogram bins are used in extracting representative edge histograms of video sequence, a global edge histogram and semi-global edge histograms for the retrieved video sequence. That is, in order to achieve a high retrieval performance, the representative edge

15   histograms having the normalized representative edge histogram bins, the global edge histogram and the semi-global edge histograms are used in the data matching process as an image descriptor for a retrieved video sequence.

20   U.S. Patent Application Serial No. 09/978,668, filed on Oct. 18th, 2001,, entitled "NON-LINEAR QUANTIZATION AND SIMILARITY MATCHING METHOD FOR RETRIEVING IMAGE DATA", commonly owned by the same assignee of this invention, the disclosure of which is incorporated by reference herein,

25   describes generation of global edge histogram and the semi-global edge histograms in detail.

In the data matching process S705, by calculating a distance between the representative, semi-global and global edge histograms of the query video sequence A and the

30   target video sequence B, a similarity between the two videos is determined as follows:

$$Distance(A,B) = \sum_{i=0}^{79} | Local\_A[i] - Local\_B[i] | + 5 \times \sum_{i=0}^{4} | Global\_A[i] - Global\_B[i] |$$
$$+ \sum_{i=0}^{64} | Semi\_Global\_A[i] - Semi\_Global\_B[i] |$$

Eq.6

35   where Local_A[i] and Local_B[i] denote, respectively,

20

index values assigned to i^{th} bins of each of representative edge histograms of the video sequences A and B; Global_A[] and Global_B[] denote, respectively, index values assigned to i^{th} bins of each of global edge histograms of the digital

5   video datas A and B; and Semi_Global_A[] and Semi_Global_B[] denote, respectively, index values assigned to i^{th} bins of each of semi-global edge histograms of the video sequences A and B. Since the number of bins for the global edge histogram is relatively smaller than that of

10   the representative and semi-global edge histograms, a weighting factor of 5 is applied in the above equation.

As explained above, using the equation 6, the similarity between the two digital video data A and B can be measured by referring to the inverse quantization tables.

15   In this case, since representative edge histogram bin values for the image should be decoded by referring to the inverse quantization tables, the equation 6 is generally used in applications for complicated but accurate retrieval. Herein, each of the inverse quantization tables is

20   corresponding to each of edge quantization tables shown in Tables 2 to 6.

The above procedure is then repeated until all of the video sequence are processed.

In accordance with the present invention, the number

25   of bits necessary for storing the quantization index values for a video sequence having a plurality of image frames can be greatly reduced. Furthermore, a complexity of the similarity calculation can be significantly decreased by using the non-linear quantizer.

30   Moreover, the present invention can be effectively retrieving digital video data including the texture video by using the edge histogram descriptor.

Although the preferred embodiments of the invention have been disclosed for illustrative purpose, those skilled

35   in the art will appreciate that various modifications, additions, and substitutions are possible, without departing from the scope and sprit of the invention as disclosed in the accompanying claims.

What is claimed is:

1. A method for constructing a database having digital video data information representing a plurality of video
5   sequence, each video sequence having a set of image frames of the digital video data, the method comprising the steps of:

a) partitioning each image frame of each video sequence into L number of sub images, wherein each sub-
10 image is further partitioned into S x T number of image-blocks, L, S and T being positive integers;

b) assigning one of 5 number of reference edges to each image-block to thereby generate L number of edge histograms for each image frame, wherein the edge
15 histograms include the M edge histogram bins and the reference edges include 4 number of directional edges and a non-directional edge;

c) normalizing the edge histogram bins contained in each edge histogram by S x T to thereby generate M number
20 of normalized edge histogram bins for said each image frame;

d) calculating M representative edge histogram bins of said each video sequence in order to generate L number of representative edge histograms of each video sequence based
25 on the normalized edge histogram bins of said each image frames; and

e) non-linearly quantizing the representative edge histogram bins to generate M number of quantization index values as a second image descriptor for said each
30 representative edge histogram to be stored in the database.

2. The method as recited in claim 1, wherein the directional edges includes a vertical edge, a horizontal edge, a 45 degree edge, a 135 degree edge and the non-
35 directional edge represents an edge of undesignated direction except for the 4 directional edges.

3. The method as recited in claim 2, wherein the

vertical edge, the horizontal edge, the 45 degree edge, the 135 degree edge and the non-directional edge are, respectively, expressed as:

vertical   horizontal   45 degree   135 degree   non-directional
edge   edge   edge   edge   edge
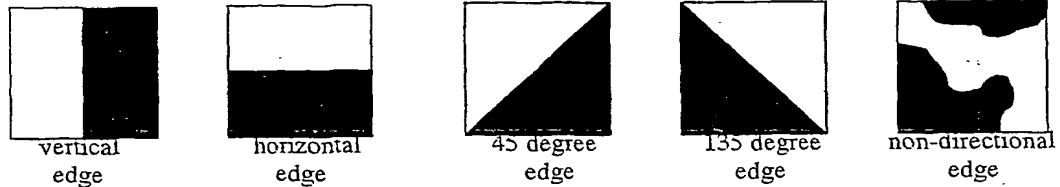
5

4. The method as recited in claim 3, wherein the step a) includes the steps of:

a-1) dividing said each image frame into N × N non-
10   overlapping sub-images to thereby form L number of rectangle-shaped sub-images, N being a positive integer; and

a-2) dividing the sub-image into S × T non-overlapping blocks to thereby form S × T number of square-shaped image-
15   blocks.

5. The method as recited in claim 4, wherein the step b) includes the steps of:

b-1) assigning one of the reference edges to each
20   image block; and

b-2) counting the number of each reference edge included in each sub-image to generate the L number of the edge histograms for the each image frame.

25   6. The method as recited in claim 5, wherein the step b-1) includes the steps of:

b-11) dividing each image-block into 2 × 2 sub-blocks;

b-12) assigning a corresponding filter coefficient to each sub-block;
30   b-13) calculating a set of 5 edge magnitudes corresponding to five edges for each image-block by using the filter coefficient; and

b-14) expressing the image-block as an edge having a maximum edge magnitude by comparing the calculated edge

magnitudes each other.

7. The method as recited in claim 6, wherein the 5 edge magnitude are obtained by using 5 number of equations, which are expressed as:

$$m_v(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_v(k) \right| ;$$

$$m_h(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_h(k) \right| ;$$

$$m_{d-45}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-45}(k) \right| ;$$

$$m_{d-135}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-135}(k) \right| ; \quad \text{and}$$

$$m_{nd}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{nd}(k) \right| , \quad \text{where} \quad m_v(i,j), \quad m_h(i,j), \quad m_{d-45}(i,j),$$

$m_{d-135}(i,j)$ and $m_{nd}(i,j)$, respectively, denote vertical, horizontal, 45 degree, 135 degree and non-directional edge magnitudes for a $(i,j)^{th}$ image-block; $a_k(i,j)$ denotes an average gray level for a sub-block assigned k in the $(i,j)^{th}$ image-block and $f_v(k)$, $f_h(k)$, $f_{d-45}(k)$, $f_{d-135}(k)$ and $f_{nd}(k)$ denote, respectively, filter coefficients for the vertical, horizontal, 45 degree, 135 degree and non-directional edges where k represents a number assigned to each sub-block.

8. The method as recited in claim 6, wherein the maximum edge magnitude is greater than a predetermined threshold value, otherwise the image block is considered to contain no edge.

9. The method as recited in claim 1, wherein the representative edge histogram bins are calculated based on a mean value of corresponding edge histogram bins of image frames of the video sequence.

10. The method as recited in claim 1, wherein the step e) includes the step of e1) non-linearly quantizing each of the representative edge histogram bins by using a

corresponding non-linear quantization table, wherein the step e1) is provided with 5 number of non-linear quantization tables, each corresponding to said each reference edge.

5

11. The method as recited in claim 10, wherein the step e1) includes the steps of:

e2-1) mapping each of the representative edge histogram bins into a representative value contained in

10   each corresponding reference edge quantization table; and

e2-2) generating 3 bits of a quantization index value representing the representative value for said each representative edge histogram bin to thereby generate L x 5 number of the quantization index values as the second image

15   descriptor for the video sequence.

12. The method as recited in claim 10, wherein the normalized edge histogram bins are non-linearly quantized by using a non-linear quantizer that is designed based on a

20   Lloyd-Max algorithm.

13. The method as recited in claim 5, wherein said N is 4.

25   14. The method as recited in claim 12, wherein the quantization table includes a table for a vertical edge histogram bin representing the number of vertical edges in the sub-image, which is expressed as:

Index	Range	Representative value
0	0.0000000 ~ 0.0343910	0.010867
1	0.0343910 ~ 0.0787205	0.057915
2	0.0787205 ~ 0.1221875	0.099526
3	0.1221875 ~ 0.1702110	0.144849
4	0.1702110 ~ 0.2280385	0.195573
5	0.2280385 ~ 0.3092675	0.260504

6	0.3092675 ~ 0.4440795	0.358031
7	0.4440795 ~ 1.0000000	0.530128

15. The method as recited in claim 12, wherein the quantization table further includes a table for a horizontal edge histogram bin representing the number of horizontal edges in the sub-image, which is expressed as:

Index	Range	Representative value
0	0.0000000 ~ 0.0411000	0.012266
1	0.0411000 ~ 0.0979065	0.069934
2	0.0979065 ~ 0.1540930	0.125879
3	0.1540930 ~ 0.2128515	0.182307
4	0.2128515 ~ 0.2789795	0.243396
5	0.2789795 ~ 0.3631455	0.314563
6	0.3631455 ~ 0.4880235	0.411728
7	0.4880235 ~ 1.0000000	0.564319

16. The method as recited in claim 12, wherein the quantization table further includes a table for a 45 degree edge histogram bin representing the number of 45 degree edges in the sub-image, which is expressed as:

Index	Range	Representative value
0	0.0000000 ~ 0.0150225	0.004193
1	0.0150255 ~ 0.0363560	0.025852
2	0.0363560 ~ 0.0576895	0.046860
3	0.0576895 ~ 0.0809025	0.068519
4	0.0809025 ~ 0.1083880	0.093286

5	0.1083880 ~ 0.1424975	0.123490
6	0.1424975 ~ 0.1952325	0.161505
7	0.1952325 ~ 1.0000000	0.228960

17. The method as recited in claim 12, wherein the quantization table further includes a table for a 135 degree edge histogram bin representing the number of 135 degree edges in the sub-image, which is expressed as:

Index	Range	Representative value
0	0.0000000 ~ 0.0150490	0.004174
1	0.0150490 ~ 0.0360780	0.025924
2	0.0360780 ~ 0.0566975	0.046232
3	0.0566975 ~ 0.0784090	0.067163
4	0.0784090 ~ 0.1025230	0.089655
5	0.1025230 ~ 0.1336475	0.115391
6	0.1336475 ~ 0.1848245	0.151904
7	0.1848245 ~ 1.0000000	0.217745

18. The method as recited in claim 12, wherein the quantization table further includes a table for a non-directional edge histogram bin representing the number of non-directional edges, which is expressed as:

Index	Range	Representative value
0	0.0000000 ~ 0.0292225	0.006778
1	0.0292225 ~ 0.0801585	0.051667
2	0.0801585 ~ 0.1374535	0.108650

3	0.1374535 ~ 0.1952415	0.166257
4	0.1952415 ~ 0.2549585	0.224226
5	0.2549585 ~ 0.3210330	0.285691
6	0.3210330 ~ 0.4036735	0.356375
7	0.4036735 ~ 1.0000000	0.450972

19.   The   method   as   recited   in   claim   7,   wherein semantics of said each edge histogram bin of BinCounts is
5 defined as:

Edge histogram bin	Semantics
BinCounts[0]	Vertical edges in sub-image (0,0)
BinCounts[1]	Horizontal edges in sub-image (0,0)
BinCounts[2]	45 degree edges in sub-image (0,0)
BinCounts[3]	135 degree edges in sub-image (0,0)
BinCounts[4]	Non-directional edges in sub-image (0,0)
BinCounts[5]	Vertical edges in sub-image (0,1)
. .	. .
BinCounts[74]	Non-directional edges in sub-image (3,2)
BinCounts[75]	Vertical edges in sub-image (3,3)
BinCounts[76]	Horizontal edges in sub-image (3,3)
BinCounts[77]	45 degree edges in sub-image (3,3)
BinCounts[78]	135 degree edges in sub-image (3,3)
BinCounts[79]	Non-directional edges in sub-image (3,3)

20.   The   method   as   recited   in   claim   1,   further comprising step f) calculating a variation value of the
10 video sequence and   storing the variation value in the database, wherein the variation value includes a variance of the edge histograms of each image frames.

21.   A   method   for   retrieving   a   corresponding   video

sequence having a set of image frames of the digital video data in response to a query video sequence based on a database, the method comprising the steps of:

     a) calculating L number of representative edge
5 histograms of the query video sequence as an image descriptor for the query video sequence, wherein each representative edge histogram represents a representative spatial distribution of 5 number of reference edges in sub-images of image frames in the query video sequence, wherein
10 the reference edges includes 4 number of directional edges and a non-directional edge;

     b) extracting a plurality of image descriptors for video sequences based on digital video data information from the database, wherein each image descriptor for said
15 each video sequence includes L number of representative edge histogram bins for said each video sequence;

     c) comparing the image descriptor for the query video sequence to said each image descriptor for each video sequences to generate a comparison result; and

20   d) retrieving at least one video sequence similar to the query video sequence based on the comparison results.

     22. The method as recited in claim 21, wherein said each edge histogram has 5 number of edge histogram bins
25 corresponding to the reference edges.

     23. The method as recited in claim 21, wherein the directional edges include a vertical edge, a horizontal edge, a 45 degree edge, a 135 degree edge and the non-
30 directional edge represents an edge of undesignated direction except for the 4 directional edges.

     24. The method as recited in claim 21, wherein the step a) includes steps of:
35   a1) partitioning each image frame of query video sequence into L number of sub images, wherein each sub-image is further partitioned into S x T number of image-blocks, L, S and T being positive integers;

a2) assigning one of 5 number of reference edges to each image-block to thereby generate L number of edge histograms for each image frame, wherein the edge histograms include the M edge histogram bins and the
5 reference edges include 4 number of directional edges and a non-directional edge;

a3) normalizing the edge histogram bins contained in each edge histogram by S x T to thereby generate M number of normalized edge histogram bins for said each image
10 frame;

a4) calculating M representative edge histogram bins of said query video sequence in order to generate L number of representative edge histograms of each video sequence based on the normalized edge histogram bins of said each
15 image frames

25. The method as recited in claim 23, wherein the step a2) includes the steps of:

a2-1) assigning one of the reference edges to each
20 image block; and

a2-2) counting the number of each reference edge included in each sub-image to generate the L number of the edge histograms for the query video sequence.

25   26. The method as recited in claim 25, wherein the step a2-1) includes the steps of:

a2-11) dividing each image-block into 2 × 2 sub-blocks;

a2-12) assigning a corresponding filter coefficient to
30 each sub-block;

a2-13) calculating a set of 5 edge magnitudes corresponding to five edges for each image-block by using the filter coefficient; and

a2-14) expressing the image-block as an edge having a
35 maximum edge magnitude by comparing the calculated edge magnitudes each other.

27. The method as recited in claim 26, wherein the 5

edge magnitude are obtained by using 5 number of equations, which are expressed as:

$$m_v(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_v(k) \right|;$$

$$m_h(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_h(k) \right|;$$

5

$$m_{d-45}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-45}(k) \right|;$$

$$m_{d-135}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{d-135}(k) \right|; \text{ and}$$

$$m_{nd}(i,j) = \left| \sum_{k=0}^{3} a_k(i,j) \times f_{nd}(k) \right|, \text{ where } m_v(i,j), \; m_h(i,j), \; m_{d-45}(i,j),$$

$m_{d-135}(i,j)$ and $m_{nd}(i,j)$, respectively, denote vertical, horizontal, 45 degree, 135 degree and non-directional edge

10 magnitudes for a $(i,j)^{th}$ image-block; $a_k(i,j)$ denotes an average gray level for a sub-block assigned k in the $(i,j)^{th}$ image-block and $f_v(k), \; f_h(k), \; f_{d-45}(k), \; f_{d-135}(k)$ and $f_{nd}(k)$ denote, respectively, filter coefficients for the vertical, horizontal, 45 degree, 135 degree and non-directional edges

15 where k represents a number assigned to each sub-block.

28. The method as recited in claim 26, wherein the maximum edge magnitude is greater than a predetermined threshold value, otherwise the image block is considered to

20 contain no edge.

29. The method as recited in claim 21, wherein the image descriptors for the query and target video sequence further include a global edge histogram and R number of the

25. semi-global histograms based on the Lx5 number of representative edge histogram bins, respectively, R being a positive integer.

30. The method as recited in claim 29, wherein the

30 global edge histogram represents an edge distribution in a whole space of the query and target video sequences and each semi-global edge histogram represents an edge

distribution in a corresponding set of sub-images of the query and target video sequences.

31. The method as recited in claim 29, wherein said N and R are 4 and 13, respectively.

32. The method as recited in claim 31, wherein each of the 13 semi-global edge histograms is generated for each of 13 sets of 4 sub-images, wherein the 13 sets include: four sets of 4 sub-images, each set including 4 sub-images in each of first to fourth columns of the image in vertical direction; four sets of 4 sub-images, each set including 4 sub-images in each of first to fourth rows in horizontal direction; four sets of 4 sub-images, each set including a corresponding sub-image and 3 sub-images neighboring the corresponding sub-image, wherein the corresponding sub-image is respectively located on the left-top, on the right-top, on the left-bottom and on the right-bottom of the image; and a set including 4 sub-images around the center of the image.

33. The method as recited in claim 21, wherein the step b) includes the steps of:

b1) retrieving L x 5 number of quantization index values for each of the target video sequence;

b2) converting each of the L x 5 number of quantization index values into L x 5 number of representative edge histogram bins for said each target vids sequence by using 5 number of non-linear inverse quantization tables; and

b3) generating L number of representative edge histograms based on the L x 5 number of normalized edge histogram bins.

34. The method as recited in claim 33, wherein the step b) further includes the step of: b4) further generating a global edge histogram and R number of semi-global histograms for each of the target video sequence

based on the L x 5 number of representative edge histogram bins.

35. The method as recited in claim 21, wherein the step b) includes the steps of:

b1) retrieving L x 5 number of quantization index values for each of the target vides sequence;

b2) converting each of the L x 5 number of quantization index values into L x 5 number of representative edge histogram bins for said each target video sequence by normalizing the L x 5 number of quantization index values; and

b3) generating L number of representative histograms based on the L x 5 number of representative edge histogram bins.

36. The method as recited in claim 35, wherein the step b) further includes the step of: b4) further generating a global edge histogram and R number of semi-global histograms for each of the target images based on the L x 5 number of normalized edge histogram bins.

37. The method as recited in claim 34, wherein the step c) includes the step of:

estimating a distance between the query video sequence and said each target video sequence by equation as:

$$Dis \tan ce(A,B) = \sum_{i=0}^{79} | Local\_A[i] - Local\_B[i] | + 5 \times \sum_{i=0}^{4} | Global\_A[i] - Global\_B[i] |$$
$$+ \sum_{i=0}^{64} | Semi\_Global\_A[i] - Semi\_Global\_B[i] |$$

where Local_A[i] and Local_B[i] denote, respectively, the edge histogram bins of BinCount[i] of the query video sequence A and the target video sequence B; Global_A[] and Global_B[] denote, respectively, the edge histogram bins for the global edge histograms of the query image A and the target image B; and Semi_Global_A[] and Semi_Global_B[]

denote, respectively, the histogram bin values for the semi-global edge histogram bins of the query video sequence A and the target video sequence B.

5   38. The method as recited in claim 36, wherein the step c) includes the step of:

estimating a distance between the query video sequence and said each target video sequence by an equation as:

$$Dis\tan ce(A,B) = \sum_{i=0}^{79} | \; Local\_A[i] - Local\_B[i] \; | + 5 \times \sum_{i=0}^{4} | \; Global\_A[i] - Global\_B[i] \; |$$

$$+ \sum_{i=0}^{64} | \; Semi\_Global\_A[i] - Semi\_Global\_B[i] \; |$$

where Local_A[i] and Local_B[i] denote, respectively, the edge histogram bins of BinCount[i] of the query video seqence A and the target video sequence B; Global_A[ ] and Global_B[ ] denote, respectively, the edge histogram bins for the global edge histograms of the query video sequence A and the target video sequence B; and Semi_Global_A[ ] and Semi_Global_B[ ] denote, respectively, the histogram bin values for the semi-global edge histogram bins of the query video sequence A and the target video sequence B.

20

39. A method for extracting an image description for video sequence, each having a plurality of image frames of the digital video data, the method comprising the steps of:

a) selecting a one of the image frames as a target image frame;

b) calculating Lx5 number of normalized edge histogram bins to generate L number of edge histograms of the target image, wherein said each edge histogram has 5 number of normalized edge histogram bins and represents a spatial distribution of 5 number of reference edges in a sub-image and L is a positive integer, wherein the reference edges include 4 number of directional edges and a non-directional edge;

c) selecting a next image frame as a target image;

d) repeating steps b) and c) until L number of edge

histograms of all image frames are calculated;

  e) calculating a representative edge histogram having LX5 number of normalized edge histogram bins for the video sequence based on the L number of edge histograms of each

5 image frame;

  f) non-linearly quantizing the Lx5 number of normalized edge histogram bins of the representative edge histograms to generate Lx5 number of quantization index values as the image description for the video sequence; and

10 g) storing the Lx5 number of quantization index values to the database.

  40. The method as recited in claim 39, wherein each quantization index value is represented by 3 bits.

15

  41. The method as recited in claim 39, wherein the directional edges include a vertical edge, a horizontal edge, a 45 degree edge, a 135 degree edge and the non-directional edge represents an edge of undesignated

20 direction except for the 4 directional edges.

25

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

A.	CLASSIFICATION OF SUBJECT MATTER

IPC7 H04N 7/24

According to International Patent Classification (IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC H04N, G06F, G06T, G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Patents and Applications for Inventions since 1975

Electronic data base consulted during the intertnational search (name of data base and, where practicable, search terms used)

NPS: "image, retrieve, quantization, edge, histogram, video, sequence, texture"

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6253201 B1 (Philips Electronics North America Corporation) 26 June 2001 See the whole document	1 - 41
A	KR 2001-0002014 A (LG Electronics Inc.) 05 January 2001 See the whole document	1 - 41
A	US 6181818 B1 (Canon Inc.) 30 January 2001 See the whole document	1 - 41

☐ Further documents are listed in the continuation of Box C.	☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01 AUGUST 2003 (01.08.2003)	01 AUGUST 2003 (01.08.2003)

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea	JEONG, Seong Jung
82-42-472-7140	Telephone No. 82-42-481-5768

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6253201 B1	26-06-2001	None	
KR 2001-0002014 A	05-01-2001	None	
US 6181818 B1	30-01-2001	JP 8137908 A2 EP 713186 A1 US 6246804 BA	31-05-1996 22-05-1996 12-06-2001